

MAR 11 2002

SEQUENCE LISTING

4

<110> Ingram, Lonnie O
Zhong, Shengde

<120> METHODS AND COMPOSITIONS FOR SIMULTANEOUS
SACCHARIFICATION AND FERMENTATION

<130> BCI-024CP

<140> 09/885,297

<141> 2001-09-19

<150> 60/214,137

<151> 2000-06-26

<150> 60/219,913

<151> 2000-07-21

<160> 17

<170> PatentIn Ver. 2.0

<210> 1

<211> 450

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:primer

<400> 1

cttttttcggc atgagcaacc aacatttttca aggtatcatc ctgatgcgca atatcggcat 60
cggttagcca taaccatttt acctgtccgg cggccttaac acctgatca gatgggttcgt 120
ggtgttggtta ccttgccgaa gggcacccgt aaaaatgttc gcgtcgggtgt tttcgcccg 180
ggcccgaaag ctgaagaagc taaagctgct ggtgcagaag ttgtcggcgc agaagacctg 240
atggaagcca ttcagggcgg cagcattgat ttcgatcgtg atgcccttta tactgaaatt 300
gccttgogct gccataatga agcagcctcc ggtgttttgg cagatttaag cgctgcctga 360
ttttcgtgat cctctagagt ctatgaaatg gagattcatt tatgcctctc tcttattcgg 420
ataaccatcc agtcatccgc aagcttgccc 450

<210> 2

<211> 1509

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:primer

<400> 2

gatcaaccgg caatttatcc acggcatcaa attcgatctg tcttttcccg tatcattggc 60
aataccggca ttctgattac aggcggtggt ttgaatggcg tatgcagttt tgtctatgtc 120

A
(pages 1-19)

#4/A

gcatggacat cccagacatt gggattgaac ctgtttggtg tcatgctttt gattacgact 180
tttgcctaccc tgatttcgga tattaccggt tttcagtcac ggcaaacctt gctgcattac 240
ggttcaaaaag cttttcagga aaaagatttt aaccaatttg atgatgtcct tgccttttgc 300
atcagagccg atttttttag tgcggcgata ggtatggttg tagggtagg cggatatctg 360
attttaggca cttcaagatt gggatggcct gccgaggtca agccagatgc cttgctttgt 420
atgctgatta tactttttat gaatatcggc tgggtccaacc gggatgttgc ggctgtgtaa 480
ccgcttttaa ctggtcacta tttatgagtt tattacgacc tgcgtcagaa ccggaggttg 540
tggcattggt tattggcttc atatgccttt ggggtatttt ttgtttatat ggtgcctgac 600
gcaattcacg ctttttgtca cctgtagtta cgctggcatt tatctctttc accaatatac 660
ggagcgagca tttccgataa gaaaaatatt tcagagaaaa acgcccgttg aagggatgtg 720
gaaattcact ttaagcgtca gttttaatga aatcctagac tccattttcc agcaggggtg 780
cacccttgct attggtagct cactgggggc tggggaagcc gctgtctatc gggcgcgcgc 840
ccagattagt aacggtttat ccaaaccagc acagatgatg atcggctaac atgcatccac 900
cggcagcacc ggccgtttta tgcttgggat tattgatatg ccgaaaagga tacaacatct 960
ggaagaaaaa gacgaaggcc ggaataagcg cccattctgc aaaattgtta caacttagtc 1020
gcgccatcag ggaatgaaaa atcaatccgt ctttttcggc atgagcaacc aacattttca 1080
aggtatcatc ctgatgcgca atatcgcat cggttagcca taaccatttt acctgtccgg 1140
cggccttaat acctgatca gatggttcgt ggtgttggtta ccttgccgaa gggcaccggt 1200
aaaaatgttc gcgtcgggtg tttcgcccggt ggcccgaaag ctgaagaagc taaagctgct 1260
ggtgcagaag ttgtcggcgc agaagacctg atggaagcca ttcagggcgg cagcattgat 1320
ttcgatcgtg atgcccttta tactgaaatt gccttgctgc gccataatga agcagcctcc 1380
ggtgttttgg cagatttaag cgctgcctga ttttcgtgat cctctagagt ctatgaaatg 1440
gagattcatt tatgcctctc tcttattcgg ataaccatcc agtcatccgc aagcttggcc 1500
gtaatccat 1509

<210> 3

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:primer

<400> 3

cgaattcctg ccgaagttta ttagcca

27

<210> 4
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 4
aaggatcctt ccaccagcta tttgttagtg a 31

<210> 5
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 5
agaattctgc cagttggttg acgatag 27

<210> 6
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 6
caggatcccc tcaagtcact agttaactg 30

<210> 7
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 7
taatacgact cactataggg 20

<210> 8
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 8
taacaatttc acacagga 18

<210> 9
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 9
cacgacgttg taaaacgac

19

<210> 10
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 10
gactggatgg ttatccgaat aagagagagg

30

<210> 11
<211> 8
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 11
ggcgcgcc

8

<210> 12
<211> 11544
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:vector

<220>
<221> misc_feature
<222> 3282-4281
<223> n=a,c,g or t

<220>
<221> misc_feature
<222> 9476-11544
<223> n=a,c,g or t

<220>
<223> nucleotide positions 1-1451 encodes promoter

<220>
<223> nucleotide positions 1452-2735 encodes celZ gene

<220>

<223> nucleotide positions 4916-5776 encodes bla gene

<220>

<223> nucleotide positions 7061-8251 encodes tet gene

<220>

<223> nucleotide positions 9476-11544 encodes target sequence from K. oxytoca

<220>

<221> CDS

<222> (1452)..(2735)

<220>

<221> CDS

<222> (4916)..(5776)

<220>

<221> CDS

<222> (7061)..(8251)

<400> 12

gatcaaccgg caatttatcc acggcatcaa attcgatctg tcttttcccg tatcattggc 60
aataccggca ttctgattac aggcctgtgt ttgaatgcgg tatgcagttt tgtctatgtc 120
gcatggacat cccagacatt gggattgaac ctgtttgggtg tcatgctttt gattacgact 180
tttgctaccc tgatttcgga tattaccogt tttcagtcac ggcaaacott gctgcattac 240
ggttcaaaag cttttcagga aaaagatttt aaccaatttg atgatgtcct tgccttttgc 300
atcagagccg attttttttag tgcggcgata ggtatgttgg tagggttagg cggatatcttg 360
atttttaggca cttcaagatt gggatggcct gccgaggtca agccagatgc cttgctttgt 420
atgctgatta tactttttat gaatatcggc tgggtccaacc gggatgttgc ggctgtgtaa 480
ccgctttaa ctggtcacta tttatgagtt tattacgacc tgcgtcagaa ccggaggttg 540
tggcattggg tattggcttc atatgccttt ggggtatttt ttgtttatat ggtgcctgac 600
gcaattcacg ctttttgtca cctgtagtta cgctggcatt tatctctttc accaatatac 660
ggagcgagca tttccgataa gaaaaatatt tcagagaaaa acgcccgttg aagggatgtg 720
gaaattcact ttaagcgtca gttttaatga aatcctagac tccattttcc agcaggggtg 780
cacccttgct attggtagct cactgggggc tggggaagcc gctgtctatc gggtcgcgcg 840
ccagattagt aacggtttat ccaaaccagc acagatgatg atcggctaac atgcatccac 900
cggcagcacc ggccgtttta tgcttgggat tattgatatg ccgaaaagga tacaacatct 960
ggaagaaaaa gacgaaggcc ggaataagcg ccattctgc aaaattgtta caacttagtc 1020
gcgccatcag ggaatgaaaa atcaatccgt ctttttcggc atgagcaacc aacattttca 1080
aggtatcatc ctgatgcgca atatcgcat cggttagcca taaccatttt acctgtccgg 1140
cggccttaat accttgatca gatgggtcgt ggtgttggtta ccttgccgaa gggcaccggg 1200

```

aaaaatgttc gcgtcggtgt ttctgcccgt ggcccgaag ctgaagaagc taaagctgct 1260
ggtgcagaag ttgtcggcgc agaagacctg atggaagcca ttcagggcgg cagcattgat 1320
ttcgatcgtg atgcccttta tactgaaatt gccttgcgct gccataatga agcagcctcc 1380
ggtgttttgg cagatttaag cgctgcctga ttttcgtgat cctctagagt ctatgaaatg 1440
gagattcatt t atg cct ctc tct tat tcg gat aac cat cca gtc atc gat 1490
          Met Pro Leu Ser Tyr Ser Asp Asn His Pro Val Ile Asp
            1              5              10

agc caa aaa cac gcc cca cgt aaa aaa ctg ttt cta tct tgt gcc tgt 1538
Ser Gln Lys His Ala Pro Arg Lys Lys Leu Phe Leu Ser Cys Ala Cys
      15              20              25

tta gga tta agc ctt gcc tgc ctt tcc agt aat gcc tgg gcg agt gtt 1586
Leu Gly Leu Ser Leu Ala Cys Leu Ser Ser Asn Ala Trp Ala Ser Val
      30              35              40              45

gag ccg tta tcc gtt agc ggc aat aaa atc tac gca ggt gaa aaa gcc 1634
Glu Pro Leu Ser Val Ser Gly Asn Lys Ile Tyr Ala Gly Glu Lys Ala
      50              55              60

aaa agt ttt gcc ggc aac agc tta ttc tgg agt aat aat ggt tgg ggt 1682
Lys Ser Phe Ala Gly Asn Ser Leu Phe Trp Ser Asn Asn Gly Trp Gly
      65              70              75

ggg gaa aaa ttc tac aca gcc gat acc gtt gcg tcg ctg aaa aaa gac 1730
Gly Glu Lys Phe Tyr Thr Ala Asp Thr Val Ala Ser Leu Lys Lys Asp
      80              85              90

tgg aaa tcc agc att gtt cgc gcc gct atg ggc gtt cag gaa agc ggt 1778
Trp Lys Ser Ser Ile Val Arg Ala Ala Met Gly Val Gln Glu Ser Gly
      95              100              105

ggt tat ctg cag gac ccg gct ggc aac aag gcc aaa gtt gaa aga gtg 1826
Gly Tyr Leu Gln Asp Pro Ala Gly Asn Lys Ala Lys Val Glu Arg Val
     110              115              120              125

gtg gat gcc gca atc gcc aac gat atg tat gtg att att gac tgg cac 1874
Val Asp Ala Ala Ile Ala Asn Asp Met Tyr Val Ile Ile Asp Trp His
      130              135              140

tca cat tct gca gaa aac aat cgc agt gaa gcc att cgc ttc ttc cag 1922
Ser His Ser Ala Glu Asn Asn Arg Ser Glu Ala Ile Arg Phe Phe Gln
      145              150              155

gaa atg gcg cgc aaa tat ggc aac aag ccg aat gtc att tat gaa atc 1970
Glu Met Ala Arg Lys Tyr Gly Asn Lys Pro Asn Val Ile Tyr Glu Ile
      160              165              170

tac aac gag ccg ctt cag gtt tca tgg agc aat acc att aaa cct tat 2018
Tyr Asn Glu Pro Leu Gln Val Ser Trp Ser Asn Thr Ile Lys Pro Tyr
      175              180              185

gcc gaa gcc gtg att tcc gcc att cgc gcc att gac ccg gat aac ctg 2066
Ala Glu Ala Val Ile Ser Ala Ile Arg Ala Ile Asp Pro Asp Asn Leu
     190              195              200              205

att att gtc ggt acg ccc agt tgg tcg caa aac gtt gat gaa gcg tcg 2114

```

Ile Ile Val Gly Thr Pro Ser Trp Ser Gln Asn Val Asp Glu Ala Ser
 210 215 220
 cgc gat cca atc aac gcc aag aat atc gcc tat acg ctg cat ttc tac 2162
 Arg Asp Pro Ile Asn Ala Lys Asn Ile Ala Tyr Thr Leu His Phe Tyr
 225 230 235
 gcg gga acc cat ggt gag tca tta cgc act aaa gcc cgc cag gcg tta 2210
 Ala Gly Thr His Gly Glu Ser Leu Arg Thr Lys Ala Arg Gln Ala Leu
 240 245 250
 aat aac ggt att gcg ctt ttc gtc acc gag tgg ggc gcc gtt aac gcg 2258
 Asn Asn Gly Ile Ala Leu Phe Val Thr Glu Trp Gly Ala Val Asn Ala
 255 260 265
 gac ggc aat ggc gga gtg aac cag aca gat acc gac gcc tgg gta acg 2306
 Asp Gly Asn Gly Gly Val Asn Gln Thr Asp Thr Asp Ala Trp Val Thr
 270 275 280 285
 ttc atg cgt gac aac aac atc agc aac gca aac tgg gcg tta aat gat 2354
 Phe Met Arg Asp Asn Asn Ile Ser Asn Ala Asn Trp Ala Leu Asn Asp
 290 295 300
 aaa agc gaa ggg gca tca acc tat tat ccg gac tct aaa aac ctg acc 2402
 Lys Ser Glu Gly Ala Ser Thr Tyr Tyr Pro Asp Ser Lys Asn Leu Thr
 305 310 315
 gag tcg ggt aaa ata gta aaa tcg atc att caa agc tgg cca tat aaa 2450
 Glu Ser Gly Lys Ile Val Lys Ser Ile Ile Gln Ser Trp Pro Tyr Lys
 320 325 330
 gcg ggc agc gcc gcc agt aca aca acc gat cag tca acc gat acc acc 2498
 Ala Gly Ser Ala Ala Ser Thr Thr Thr Asp Gln Ser Thr Asp Thr Thr
 335 340 345
 atg gca cca ccg ttg acg aac cga cca caa ccg aca cac cgg caa acc 2546
 Met Ala Pro Pro Leu Thr Asn Arg Pro Gln Pro Thr His Arg Gln Thr
 350 355 360 365
 gct gat tgc tgc aat gcc aac gtt tac ccc aac tgg gtt agc aaa gac 2594
 Ala Asp Cys Cys Asn Ala Asn Val Tyr Pro Asn Trp Val Ser Lys Asp
 370 375 380
 tgg gcg ggc cgg cag cga ctc ata acg aag cag gcc aat cga tcg tct 2642
 Trp Ala Gly Arg Gln Arg Leu Ile Thr Lys Gln Ala Asn Arg Ser Ser
 385 390 395
 aca aag gga acc tgt ata ccg caa act ggt aca ctt cat ccg ttc cgg 2690
 Thr Lys Gly Thr Cys Ile Pro Gln Thr Gly Thr Leu His Pro Phe Arg
 400 405 410
 gca gcg att cct cct ggg cac agg ttg gta gct gta act aat tga 2735
 Ala Ala Ile Pro Pro Gly His Arg Leu Val Ala Val Thr Asn
 415 420 425
 ttaatctttt caccgccaaa ataacagggc tgcgattgca gcctgatacg caacattcca 2795
 ttacttaatt gcgttcaaaa gcgcccaaat ccggtgcgct gccttgtaac taatatgatt 2855
 tctctttcgt acccgcttta atcagctttg agttagccga cagacggaac agcgaggttg 2915

ccggcaacgt gccgtcatta tcacgagata cggtagccag cgaggtgtcc aggctgacga 2975
atcggacgcg gaagccgctg tccgtatcca tgagttgact cgcacccgca ttactgaccg 3035
ttgcagaagc agacagagac acgttggtgc ggaagtaatg tttctgtcct gactggacgt 3095
tgctcccgaagc agcataatta atgccgtttt tatatgacgt gttatttatt accgtacgcc 3155
gccgcgttat tgttctggtc aaaacctttg ctcacgttgc caaacgcgcac gcaacgggta 3215
atgcgatgat tgccgaccgc tggttcctcc cagtttgaac ccgttgcat tgccggcgaa 3275
cgcgctnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3335
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3395
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3455
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3515
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3575
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3635
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3695
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3755
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3815
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3875
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3935
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3995
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4055
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4115
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4175
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 4235
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnngatc ctctagagtc 4295
gacctgcagg aattcactgg ccgtcgtttt acaacgtcgt gactgggaaa accctggcgt 4355
taccgaactt aatcgccctg cagcacatcc ccctttcgcc agctggcgta atagcgaaga 4415
ggcccgacc gatcgccctt cccaacagtt gcgcagcctg aatggcgaat ggcgcctgat 4475
gcggtatttt ctcttacgc atctgtgcgg tatttcacac cgcataggcg cgcctatggg 4535
gcactctcag tacaatctgc tctgatgccg catagttaag ccagccccga caccgcgcaa 4595
caccgctga cgcgcctga cgggcttgc tgctcccggc atccgcttac agacaagctg 4655
tgaccgtctc cgggagctgc atgtgtcaga ggttttcacc gtcacaccg aaacgcgcga 4715
gacgaaaggg cctcgtgata cgcctatttt tatagggttaa tgtcatgata ataatggttt 4775

cttagacgctc aggtggcact tttcggggaa atgtgcgcgg aaccacctatt tgtttatttt	4835
tctaaataca ttcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat	4895
aatattgaaa aaggaagagt atg agt att caa cat ttc cgt gtc gcc ctt att	4948
Met Ser Ile Gln His Phe Arg Val Ala Leu Ile	
1 5 10	
ccc ttt ttt gcg gca ttt tgc ctt cct gtt ttt gct cac cca gaa acg	4996
Pro Phe Phe Ala Ala Phe Cys Leu Pro Val Phe Ala His Pro Glu Thr	
15 20 25	
ctg gtg aaa gta aaa gat gct gaa gat cag ttg ggt gca cga gtg ggt	5044
Leu Val Lys Val Lys Asp Ala Glu Asp Gln Leu Gly Ala Arg Val Gly	
30 35 40	
tac atc gaa ctg gat ctc aac agc ggt aag atc ctt gag agt ttt cgc	5092
Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys Ile Leu Glu Ser Phe Arg	
45 50 55	
ccc gaa gaa cgt ttt cca atg atg agc act ttt aaa gtt ctg cta tgt	5140
Pro Glu Glu Arg Phe Pro Met Met Ser Thr Phe Lys Val Leu Leu Cys	
60 65 70 75	
ggc gcg gta tta tcc cgt att gac gcc ggg caa gag caa ctc ggt cgc	5188
Gly Ala Val Leu Ser Arg Ile Asp Ala Gly Gln Glu Gln Leu Gly Arg	
80 85 90	
cgc ata cac tat tct cag aat gac ttg gtt gag tac tca cca gtc aca	5236
Arg Ile His Tyr Ser Gln Asn Asp Leu Val Glu Tyr Ser Pro Val Thr	
95 100 105	
gaa aag cat ctt acg gat ggc atg aca gta aga gaa tta tgc agt gct	5284
Glu Lys His Leu Thr Asp Gly Met Thr Val Arg Glu Leu Cys Ser Ala	
110 115 120	
gcc ata acc atg agt gat aac act gcg gcc aac tta ctt ctg aca acg	5332
Ala Ile Thr Met Ser Asp Asn Thr Ala Ala Asn Leu Leu Leu Thr Thr	
125 130 135	
atc gga gga ccg aag gag cta acc gct ttt ttg cac aac atg ggg gat	5380
Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp	
140 145 150 155	
cat gta act cgc ctt gat cgt tgg gaa ccg gag ctg aat gaa gcc ata	5428
His Val Thr Arg Leu Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala Ile	
160 165 170	
cca aac gac gag cgt gac acc acg atg cct gta gca atg gca aca acg	5476
Pro Asn Asp Glu Arg Asp Thr Thr Met Pro Val Ala Met Ala Thr Thr	
175 180 185	
ttg cgc aaa cta tta act ggc gaa cta ctt act cta gct tcc cgg caa	5524
Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu Thr Leu Ala Ser Arg Gln	
190 200	
caa tta ata gac tgg atg gag gcg gat aaa gtt gca gga cca ctt ctg	5572
Gln Leu Ile Asp Trp Met Glu Ala Asp Lys Val Ala Gly Pro Leu Leu	
205 210 215	

cgc tcg gcc ctt ccg gct ggc tgg ttt att gct gat aaa tct gga gcc 5620
 Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile Ala Asp Lys Ser Gly Ala
 220 225 230 235

ggt gag cgt ggg tct cgc ggt atc att gca gca ctg ggg cca gat ggt 5668
 Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala Ala Leu Gly Pro Asp Gly
 240 245 250

aag ccc tcc cgt atc gta gtt atc tac acg acg ggg agt cag gca act 5716
 Lys Pro Ser Arg Ile Val Val Ile Tyr Thr Thr Gly Ser Gln Ala Thr
 255 260 265

atg gat gaa cga aat aga cag atc gct gag ata ggt gcc tca ctg att 5764
 Met Asp Glu Arg Asn Arg Gln Ile Ala Glu Ile Gly Ala Ser Leu Ile
 270 275 280

aag cat tgg taa ctgtcagacc aagtttactc atatatactt tagattgatt 5816
 Lys His Trp
 285

taaaacttca tttttaattt aaaaggatct aggtgaagat cctttttgat aatctcatga 5876

ccaaaatccc ttaacgtgag ttttcgttcc actgagcgtc agaccccgta gaaaagatca 5936

aaggatcttc ttgagatcct ttttttctgc gcgtaatctg ctgcttgcaa acaaaaaaac 5996

caccgctacc agcgggtggtt tgtttgccgg atcaagagct accaactctt tttccgaagg 6056

taactggctt cagcagagcg cagataccaa atactgtcct tctagtgtag ccgtagttag 6116

gccaccactt caagaactct gtagcaccgc ctacatacct cgctctgcta atcctgttac 6176

cagtggctgc tgccagtggc gataagtcgt gtcttaccgg gttggactca agacgatagt 6236

taccggataa ggcgagcggg tcgggctgaa cgggggggtc gtgcacacag ccagcttg 6296

agcgaacgac ctacaccgaa ctgagatacc tacagcgtga gctatgagaa agcgccacgc 6356

ttcccgaagg gagaaaggcg gacaggtatc cggtaagcgg cagggtcgga acaggagagc 6416

gcacgagggg gcttcagggg ggaaacgcct ggtatcttta tagtcctgtc gggtttcgcc 6476

acctctgact tgagcgtcga tttttgtgat gctcgtcagg ggggaggagc ctatggaaaa 6536

acgccagcaa cgcggccttt ttacggttcc tggccttttg ctggcctttt gctcacatgt 6596

tctttcctgc gttatcccct gattctgtgg ataaccgtat taccgccttt gagttagctg 6656

ataccgctcg ccgagccga acgaccgagc gcagcgagtc agtgagcgag gaagcggcgc 6716

gccagcgga gagcgcccaa tacgcaaacc gcctctccc gcgcgttggc cgattcatta 6776

atgcagctgg cagcagaggt ttcccgaact gaaagcgggc agtgagcgca acgcaattaa 6836

tgtgagttag ctactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat 6896

gttgtgtgga attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta 6956

cgccaagctt gcatgccaat tctcatgttt gacagcttat catcgataag ctttaaatgcg 7016

gtagtttatc acagttaaat tgctaacgca gtcaggcacc gtgt atg aaa tct aac 7072

Met Lys Ser Asn

aat gcg ctc atc gtc atc ctc ggc acc gtc acc ctg gat gct gta ggc	7120
Asn Ala Leu Ile Val Ile Leu Gly Thr Val Thr Leu Asp Ala Val Gly	
5 10 15 20	
ata ggc ttg gtt atg ccg gta ctg ccg ggc ctc ttg cgg gat atc gtc	7168
Ile Gly Leu Val Met Pro Val Leu Pro Gly Leu Leu Arg Asp Ile Val	
25 30 35	
cat tcc gac agc atc gcc agt cac tat ggc gtg ctg cta gcg cta tat	7216
His Ser Asp Ser Ile Ala Ser His Tyr Gly Val Leu Leu Ala Leu Tyr	
40 45 50	
gcg ttg atg caa ttt cta tgc gca ccc gtt ctc gga gca ctg tcc gac	7264
Ala Leu Met Gln Phe Leu Cys Ala Pro Val Leu Gly Ala Leu Ser Asp	
55 60 65	
cgc ttt ggc cgc cgc cca gtc ctg ctc gct tcg cta ctt gga gcc act	7312
Arg Phe Gly Arg Arg Pro Val Leu Leu Ala Ser Leu Leu Gly Ala Thr	
70 75 80	
atc gac tac gcg atc atg gcg acc aca ccc gtc ctg tgg atc ctc tac	7360
Ile Asp Tyr Ala Ile Met Ala Thr Thr Pro Val Leu Trp Ile Leu Tyr	
85 90 95 100	
gcc gga cgc atc gtg gcc gcc atc acc ggc gcc aca ggt gcg gtt gct	7408
Ala Gly Arg Ile Val Ala Gly Ile Thr Gly Ala Thr Gly Ala Val Ala	
105 110 115	
ggc gcc tat atc gcc gac atc acc gat ggg gaa gat cgg gct cgc cac	7456
Gly Ala Tyr Ile Ala Asp Ile Thr Asp Gly Glu Asp Arg Ala Arg His	
120 125 130	
ttc ggg ctc atg agc gct tgt ttc ggc gtg ggt atg gtg gca ggc ccc	7504
Phe Gly Leu Met Ser Ala Cys Phe Gly Val Gly Met Val Ala Gly Pro	
135 140 145	
gtg gcc ggg gga ctg ttg gcc gcc atc tcc ttg cat gca cca ttc ctt	7552
Val Ala Gly Gly Leu Leu Gly Ala Ile Ser Leu His Ala Pro Phe Leu	
150 155 160	
gcg gcg gcg gtg ctc aac ggc ctc aac cta cta ctg ggc tgc ttc cta	7600
Ala Ala Ala Val Leu Asn Gly Leu Asn Leu Leu Leu Gly Cys Phe Leu	
165 170 175 180	
atg cag gag tcg cat aag gga gag cgt cga ccg atg ccc ttg aga gcc	7648
Met Gln Glu Ser His Lys Gly Glu Arg Arg Pro Met Pro Leu Arg Ala	
185 190 195	
ttc aac cca gtc agc tcc ttc cgg tgg gcg cgg ggc atg act atc gtc	7696
Phe Asn Pro Val Ser Ser Phe Arg Trp Ala Arg Gly Met Thr Ile Val	
200 205 210	
gcc gca ctt atg act gtc ttc ttt atc atg caa ctc gta gga cag gtg	7744
Ala Ala Leu Met Thr Val Phe Phe Ile Met Gln Leu Val Gly Gln Val	
215 220 225	
ccg gca gcg ctc tgg gtc att ttc ggc gag gac cgc ttt cgc tgg agc	7792
Pro Ala Ala Leu Trp Val Ile Phe Gly Glu Asp Arg Phe Arg Trp Ser	
230 235 240	

gcg acg atg atc ggc ctg tcg ctt gcg gta ttc gga atc ttg cac gcc 7840
 Ala Thr Met Ile Gly Leu Ser Leu Ala Val Phe Gly Ile Leu His Ala
 245 250 255 260

ctc gct caa gcc ttc gtc act ggt ccc gcc acc aaa cgt ttc ggc gag 7888
 Leu Ala Gln Ala Phe Val Thr Gly Pro Ala Thr Lys Arg Phe Gly Glu
 265 270 275

aag cag gcc att atc gcc ggc atg gcg gcc gac gcg ctg ggc tac gtc 7936
 Lys Gln Ala Ile Ile Ala Gly Met Ala Ala Asp Ala Leu Gly Tyr Val
 280 285 290

ttg ctg gcg ttc gcg acg cga ggc tgg atg gcc ttc ccc att atg att 7984
 Leu Leu Ala Phe Ala Thr Arg Gly Trp Met Ala Phe Pro Ile Met Ile
 295 300 305

ctt ctc gct tcc ggc ggc atc ggg atg ccc gcg ttg cag gcc atg ctg 8032
 Leu Leu Ala Ser Gly Gly Ile Gly Met Pro Ala Leu Gln Ala Met Leu
 310 315 320

tcc agg cag gta gat gac gac cat cag gga cag ctt caa gga tcg ctc 8080
 Ser Arg Gln Val Asp Asp Asp His Gln Gly Gln Leu Gln Gly Ser Leu
 325 330 335 340

gcg gct ctt acc agc cta act tcg atc act gga ccg ctg atc gtc acg 8128
 Ala Ala Leu Thr Ser Leu Thr Ser Ile Thr Gly Pro Leu Ile Val Thr
 345 350 355

gcg att tat gcc gcc tcg gcg agc aca tgg aac ggg ttg gca tgg att 8176
 Ala Ile Tyr Ala Ala Ser Ala Ser Thr Trp Asn Gly Leu Ala Trp Ile
 360 365 370

gta ggc gcc gcc cta tac ctt gtc tgc ctc ccc gcg ttg cgt cgc ggt 8224
 Val Gly Ala Ala Leu Tyr Leu Val Cys Leu Pro Ala Leu Arg Arg Gly
 375 380 385

gca tgg agc cgg gcc acc tcg acc tga atggaagccg gcggcacctc 8271
 Ala Trp Ser Arg Ala Thr Ser Thr
 390 395

gctaacggat tcaccactcc aagaattgga gccaatcaat tcttgccgag aactgtgaat 8331

gcgcaaacca acccttggca gaacatatcc atcgcgctccg ccatctccag cagccgcacg 8391

cggcgcatct cggggctcgac tctagaggat ccccgcaacg ctgtcagcgc tttccagtta 8451

aacggctcca acgtcgccat aggtaattcc tcgcccggcc atacgatcgg gcaggtgccg 8511

ttggctatcg ccgtcgctg actcatcaca ctatcttccg ctgcatcggc aagggttttg 8571

accacttctt ccatctctcc gtgcgcggga tgccatgctc acgtacgcgg cttatcagat 8631

agtcgggcag gccgtcgttc cagcccaatg aggggaagct ggcgtggagc gatgccagca 8691

cctgctcctc aacaccgtaa tggccggcgg cgaacaggca ttccggcgta agcgcttcca 8751

gccctttaat catcacgtg cggcacatct tgatagccga cacgctgcc aagtggttac 8811

caccatagcg ggcgttacat ccaagcgtgg tgagtaattc agcaattgcc tctgcctgtg 8871

gtccccccgt caacagcggc gttcggagtg cccctggggg gaccggcgcc atcaccgcta 8931
catcgacata agcgccgggc ttaaagcatt tggcagcctg acgcttggtc tgcggggcga 8991
cggagttaag gtcaagaaaa tactgctgtg cggtcatcag cgggtgcagct tgtgaggcga 9051
catccagggc ggatcccgcg gtgacggtgg aaaatatgag ttcggcacct gtcaacgcgt 9111
cagccaggga gattgccgcc cgcacgcctc cccgatgcgc cttcgttata atcgcatcgc 9171
gctcaggacc ttgcagcttg caatcccaga cggtgactgg gttcactttt gccagtgcac 9231
ccgcaagaat gccacctgct tcaccataac ctataaacgt tattgtcgtc ataacagctc 9291
cttacgcggc cacacgtcgg ccggaatgca aacgtcgccc gcgaacagaa gtcgcgccgt 9351
acgcagcaga ccgcagcctg ccaactgccc attatcatca agccggagcg ccacgctgaa 9411
ttgggtaccg agctccgaat tgggtaccga gctcgaatta attcgagctc ggtaccgggg 9471
gatcnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 9531
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 9591
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 9651
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 9711
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 9771
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 9831
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 9891
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 9951
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10011
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10071
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10131
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10191
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10251
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10311
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10371
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10431
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10491
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10551
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10611
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10671
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 10731

nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 10791
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 10851
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 10911
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 10971
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 11031
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 11091
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 11151
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 11211
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 11271
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 11331
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 11391
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 11451
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 11511
 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 11544

<210> 13
 <211> 18
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Primer

<400> 13
 ctgttcggtt accaacac 18

<210> 14
 <211> 18
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Primer

<400> 14
 gtgaatggga tcacgagt 18

<210> 15
 <211> 26
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Primer

<400> 15
accatcagca tcaacgccca acaacg

26

<210> 16
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 16
gactggatgg ttatccgaat aagagagagg

30

<210> 17
<211> 11772
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> 134-1143
<223> n=a,c,g or t

<220>
<221> misc_feature
<222> 5753-7567
<223> n=a,c,g or t

<220>
<223> Description of Artificial Sequence: Synthetic
construct

<220>
<223> unsequenced Erwinia DNA fragment from nucleotide
position 144 to 1143

<220>
<223> P1 promoter region for celZ from nucleotide
position 2974 to 4424

<220>
<223> guide fragment for integration from nucleotide
position 4677 to 7573

<220>
<223> sequenced partial guide fragment from nucleotide
position 4677 to 5752

<220>
<223> unsequenced partial guide fragment from nucleotide
position 5753 to 7573

<220>
<223> P2 promoter region for celY from nucleotide
position 7585 to 8576

<220>
<223> R6K-Y ori from nucleotide position 10388 to 10763

<220>

<223> FRTF lipase-binding sequence from nucleotide
position 16 to 50

<220>

<223> FRTFlipase-binding sequence from nucleotide
position 10058 to 10092

<220>

<223> celZ gene product is encoded by the complement of
nucleotides 1690 to 2973

<220>

<223> celY gene product is encoded by the nucleotides
8576 to 9574

<220>

<223> kanamycin-resistance gene product is encoded by
the complement of nucleotides 10827 to 11621

<400> 17

```
atcgatgaat tgatccgaag ttcctattct ctagaaagta taggaacttc gaattgtcga 60
caagcttgat ctggcttatc gaaattaata cgactcacta tagggagacc ggaattcccc 120
tgcaggtcga ctctagagga tcannnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 180
nnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 240
nnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 300
nnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 360
nnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 420
nnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 480
nnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 540
nnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 600
nnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 660
nnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 720
nnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 780
nnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 840
nnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 900
nnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 960
nnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 1020
nnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 1080
nnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 1140
nnnagcgcgt tcgccggcaa tgccaacggg ttcaaaactgg gaggaaccag cggtcggcaa 1200
tcacgcgcat acccggttgcg tcgcgtttgg caacgtgagc aaagggttttg accagaacaa 1260
taacgcggcg gcgtacggta ataaataaca cgtcatataa aaacggcatt aattatgctt 1320
tcgggagcaa cgtccagtea ggacagaaac attacttcg caacaacgtg tctctgtctg 1380
cttctgcaac ggtcagtaat gcggatgcga gtcaactcat ggatacggac agcggcttcc 1440
gcgtccgatt cgtcagcctg gacacctcgc tggctaccgt atctcgtgat aatgacggca 1500
cggttgcggc aacctcgtg ttccgtctgt cggctaactc aaagctgatt aacgcgggta 1560
cgaaagagaa atcatattag ttacaaggca gcgcaccgga tttgggcgct tttgaacgca 1620
attaagtaat ggaatgttgc gtatcaggct gcaatcgag ccctgttatt ttgggggtga 1680
aaagattaat caattagtta cagctaccaa cctgtgccca ggaggaatcg ctgcccgaa 1740
cggatgaagt gtaccagttt gcggtataca ggttcccttt gtagacgatc gattggcctg 1800
cttcgttatg agtcgctgcc ggcccggcca gtctttgcta acccagttgg ggtaaactgt 1860
ggcattgcag caatcagcgg tttgccggtg tgctcggttg ggtcggttcg tcaacgggtg 1920
tgccatggtg gtatcggttg actgatcggg ttgtgtactg gcggcgctgc ccgctttata 1980
tgccagctt tgaatgatcg attttactat tttacccgac tcggtcagggt ttttagagtc 2040
cggataatag gttgatgcc cttcgttttt atcatttaac gccagtttg cgttgctgat 2100
gttgttgtca cgcataagc ttaccaggc gtcggtatct gtctggttca ctccgccatt 2160
gccgtccgcg ttaacggcgc cccactcggg gacgaaaagc gcaataccgt tatttaacgc 2220
ctggcgggct ttagtgcgta atgactcacc atgggttccc gcgtagaaat gcagcgtata 2280
ggcgatattc ttggcggtga ttggatcgcg cgacgcttca tcaacgtttt gcgaccaact 2340
ggcggtaccg acaataatca ggttatccgg gtcaatggcg cgaatggcgg aaatcacggc 2400
ttcggcataa ggtttaatgg tattgtctca tgaaacctga agcggctcgt tgtagatttc 2460
```


ataaatgaca ttccgcttgt tgccatatatt gcgcgccatt tccctggaaga agcgaatggc 2520
ttcactgcga ttgttttctg cagaatgtga gtgccagtca ataatcacat acatatcggt 2580
ggcgattgcg gcatccacca ctctttcaac tttggccttg ttgccagccg ggtcctgcag 2640
ataaccaccg ctttcctgaa cgcccatagc ggcgcgaaaca atgctggatt tccagtcctt 2700
tttcagcgac gcaacggtat cggctgtgta gaatttttcc ccaccccaac cattattact 2760
ccagaataag ctgttgccgg caaaactttt ggctttttca cctgcgtaga ttttattgcc 2820
gctaaccgat aacggctcaa cactcgccca ggcattactg gaaaggcagg caaggcttaa 2880
tcctaaacag gcacaagata gaaacagttt tttacgtggg gcgtgttttt ggctatcgat 2940
gactggatgg ttatccgaat aagagagagg cataaatgaa tctccatttc atagactcta 3000
gaggatcacg aaaatcaggc agcgcttaaa tctgccaaaa caccggaggc tgcttcatta 3060
tggcagcgca aggcaatttc agtataaagg gcatcacgat cgaaatcaat gctgccgccc 3120
tgaatggctt ccatcagggtc ttctgcgccg acaacttctg caccagcagc tttagcttct 3180
tcagctttcg ggccacgggc gaaaacaccg acgcgaacat ttttaccggt gcccttcggc 3240
aaggtaaaca caccacgaac catctgatca aggtattaag gccgccggac aggtaaaatg 3300
gttatggcta accgatggcg atattgcgca tcaggatgat acctgaaaa tgttggttcg 3360
tcatgccgaa aaagacggat tgatttttca ttccctgatg gcgcgactaa gttgtaaaca 3420
ttttgcagaa tgggcgctta ttccggcctt cgtctttttc ttccagatgt tgtatccttt 3480
tcggcatatc aataatccca agcataaaac ggccggtgct gccggtggat gcatgttagc 3540
cgatcatcat ctgtgctggg ttggataaac cgttactaat ctggcgcgcg acccgataga 3600
cagcggtctc cccagccccc agtgagctac caatagcaag ggtgccaccc tgctggaaaa 3660
tggagtctag gatttcatta aaactgacgc ttaaagttaa tttccacatc ccttcaacgg 3720
gcgtttttct ctgaaatatt tttcttatcg gaaatgctcg ctccgtatat tggtgaaaa 3780
gataaatgcc agcgtaacta caggtgacaa aaagcgtgaa ttgcgtcagg caccatataa 3840
acaaaaata ccccaaaggc atatgaagca aataaccaat gccacaacct ccggttctga 3900
cgagggtcgt aataaactca taaatagtga ccagtttaaa gcggttacac agccgcaaca 3960
tcccggttgg accagccgat attcataaaa agtataatca gcatacaaa caaggcatct 4020
ggcttgacct cggcaggcca tcccaatctt gaagtgccta aaatcaagat accgcctaac 4080
cctaccaaca tacctatcgc cgactaaaa aaatcggtc tgatgcaaaa ggcaaggaca 4140
tcatcaaat ggtaaaaatc tttttcctga aaagcttttg aaccgtaatg cagcaagggt 4200
tgccatgact gaaaacgggt aatatccgaa atcagggtag caaaagtcgt aatcaaaaagc 4260
atgacacca acaggttcaa tcccaatgct tgggatgtcc atgcgacata gacaaaactg 4320
cataccgat tcaaaacag gcctgtaate agaatgccgg tattgccaat gatacgggaa 4380
aagacgatc gaatttgatg ccgtggataa attgccggtt gatcgatccc cgggtaccga 4440
gctcgaattc cgagcttggc gcgcctatgc ggtgtgaaat accgcacaga tgcgtaaaga 4500
gaaaataacc catcaggcgc cattcgccat tcaggctgcg caactgttgg gaaggcgcat 4560
cgggtgcggc ctcttcgcta ttacgccagc tggcgaaagg gggatgtgct gcaaggcgat 4620
taagttgggt aacgccaggg ttttcccagt cacgacgttg taaaacgacg gccagtgaat 4680
tccatcaacg cttgctgtaa ccaggagcca aagctatgaa tgtaaccttt agctactcgt 4740
caccacacct gagcggttag gcgttaaagc actctattgc ttataagctg atgtttatca 4800
tcggcaaaaga cccggctatc gctaacaagc atgaatggct caacgccacg ctgttcgccc 4860
ttcgcatcg tatggttagg cgctggctgc gctcaaaccg cgcgcacgtc tctcaggaa 4920
ttcgccagg ttactacctg tcgatggaat ttttgattgg ccgtacgttg tccaacgcgc 4980
tgttatcgct cggcatttat gaggatgtga acagcgcgct ggaagagatg gggctgaacc 5040
ttgaagaatt aattgatgaa gaaaacgacc cgggcttagg caacggcggt cttggtcgtc 5100
tggcggcctg ctctctcgat tcgcttgccg cgctggggtt accgggcccgc ggctacggt 5160
ttcgctacga ctacgggatg tttaagcaga atatcgtcga tgggcggcag aaagaatccc 5220
cggattactg gctggaatac ggtaaccctg gggagttcga gcgccataat acgcgctaca 5280
aagtgcgctt cggcggacgc attcagcagg aaggtaaata ctcccgtgg gtggagaccg 5340
aagagattat tgccgaagcc tatgaccaga ttatccctgg cttcgacacc gacgccacca 5400
acacgctgcg cctgtggagc gcccaggcca gcagcgagat taacctcgg aaattcaacc 5460
aggcgacta cttcgcggcg gtggaagata aaaaccattc cgagaacgtg tcgcgggtac 5520
totatccgga tgactcgacc tattcaggac gcgagctgcg cctgcggcag gactacttc 5580
tcgtttcggc gacggtgcag gacatcctca gccgccacta ccagctgcat aaaacctacg 5640
ccaacctggc ggacaaaatc gcgattcatc tcaacgacac gaaccgggtg ctgtcgattc 5700
cggagctgat gcgcctgctg attgacgagc ataagatcag ctgggatgag ggnnnnnnnn 5760
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 5820
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 5880
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 5940
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6000
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6060
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 6120

nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 6180
nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 6240
nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 6300
nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 6360
nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 6420
nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 6480
nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 6540
nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 6600
nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 6660
nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 6720
nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 6780
nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 6840
nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 6900
nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 6960
nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 7020
nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 7080
nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 7140
nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 7200
nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 7260
nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 7320
nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 7380
nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 7440
nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 7500
nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn 7560
nnnnnnnnncaa ttcgagctcg gtaccgagg atcgacataa ccgataggctc ctgcattgat 7620
ggactgaaag gtttcgacgg cggttggtgt ggttttgctt gcccatctgg cggcatgaat 7680
agtgtcattc atgacgatcc agttcgatat tcaacagacc gtctttgtaa tcggcaccga 7740
caattttgat caataaagcg tttgacctga tgcgatgagg taaatccatt cgttcgggtg 7800
ttcttttctg attacctgtc ctgttaacct gtggatatag aaggctcgggt caatgagtag 7860
tattctgacg catctgacaa ttggttccaa tgacctgaag aaggcgcgca tcttttatga 7920
tgctgttttg gaaccgttg gtatcaaact tattcgcgag gtcgaaggac agcgttttgc 7980
ctatggtaaa gacggcgaag aaggacgcac catcattgta aagcctatta atggtgaagc 8040
cgctaccgct ggaaatggta tcaactatcg tttggcagcg ccttctgatg aagctgtcga 8100
tgctttttat aaagcaggct tggctaattg cggttaaggat gccggagaaac cggggcctcg 8160
tccggtctgt aataattctc ggggtgccta tttatatgac cctgaaggca ataaaatctg 8220
cgctttcaat tttaaataag atttcttttg tgcagggtta ttcaaaatag ccctgcattt 8280
tcagtattat agcggccatt atggcttttg ccttgataaa aaatttatca gggctgtttt 8340
tcgtgatgaa tatttttgat ttttcaagaa aagcctgata tcttccaaca tctttctttg 8400
tatataaatg gagcgagcta tggcgcgctg aactgtcgaa gactgtatcg ataaagttca 8460
taatcgtttc gatttgatcc tctagagtca acctgttgt tactcgtgat ccattcaca 8520
aggcgcaatt aattcgccct tctgttccgt taccaacact tgagccggag gcataatggg 8580
aaagccaatg tggcgttgtt gggcgttgat gctgatggtg tggttcagtg cgtcggctac 8640
ggcgcgaaag ggctgggaaa tctataaaag ccgtttcatg accacggacg ggcgcatcca 8700
ggataccggc aataagaatg tcagccacac cgaaggtcag ggattcgcca tgctgatggc 8760
ggtgcattac gatgacgca tcgcgttcga taacctgtgg aactggacgc aaagccacct 8820
gcggaacacg accagcggt tgttctactg gcgttacgat ccgtcggcgg ccaatccggt 8880
ggtggataag aacaacgcct cggtatggcg tgtgctgatt gcctggcggt tgttaaaagc 8940
gggaaataag tggcaggaca accgttacct gcaggcgctg gacagcatcc agaaagcgat 9000
catcgccagc aatatcattc agtttgcggt ccgcaccgtg atgttgcccg gcgcctatgg 9060
tttcaacaag aacagctatg tgatccttaa cccgtcgtat tctctgttcc cggcctggcg 9120
cgactttgct aaccgcagcc atcttcagggt gtggcgcaa ctgattgacg acagcctgtc 9180
attggctcga gaaatgcgtt tcggtcagggt cgggctgcgg acggactggg cggcgctgaa 9240
cgcggtatgg tcgatggcg cggcgacggc ctggcgctcg cgtttcagtt acgacgcat 9300
tcgtatcccg ctgtatttgt actggtatga cgccaaaacc acggcgctgg tgccgttcca 9360
gctgtactgg cgtaactatc cccgcctgac gacgcggcc tgggttgatg tgctgagcag 9420
taacaccgcg acttacaata tgcaggcggt tttgctggcg gtgcgcgacc tgacgatggg 9480
caacctcgac gggctcagcg atctgccagg cgcacggaa gattactact cgtcgagcct 9540
gcgcctgctg gtgatgttg cgcgcgtaa ataaccttat tcttgcgga cacatggcga 9600
ggacgatgtc cttgccattt tcccacttt tatcctctg aatggcggt ttttcacgct 9660
ttgttaacct cttgtttact cgtgatccca ttcacaagg cgaattgacc tgcaggcatg 9720
caagcttggc gtaatcatgg tcatagctgt ttctgtgtg aaattgttat ccgctcaca 9780

ttccacacaa	catacgagcc	ggaagcataa	agtgtaaagc	ctgggggtgcc	taatgagtga	9840
gctaactcac	attaattgcg	ttgcgctcac	tgcccgcctt	ccagtcggga	aacctgtcgt	9900
gccagctgca	ttaatgaatc	ggccaacgcg	cggggagagg	cggtttgct	attgggctgt	9960
cttcgcgtgg	cgcgccaggt	cgactctaga	ggatccccgg	ggaagatctt	ccggaagatc	10020
ttcccagagct	cgaattaatt	ccgcgatgaa	ttgatccccg	aagttcctat	tctctagaaa	10080
gtataggaac	tcgaattggt	cgacaagcta	gcttgcatgc	aagcttgat	tctatagtgt	10140
cacctaaatc	gtatgtgtat	gatacataag	gttatgtatt	aattgtagcc	gcgttctaac	10200
gacaatatgt	acaagcctaa	ttgtgtagca	tctgggttac	tgaagcagac	cctatcatct	10260
ctctcgtaaa	ctgccgtcag	agtcggtttg	gttggaacga	ccttctgagt	ttctggtaac	10320
gccgttccgc	accccgga	tggtcagcga	accaatcagc	agggtcacgc	ctagcccatg	10380
gctaattctg	tcagccgtta	agtgttcctg	tgctactgaa	aattgctttg	agaggctcta	10440
agggcttctc	agtgcgttac	atccctggct	tgttggtccac	aaccgttaaa	ccttaaaagc	10500
tttaaaagcc	ttatatattc	ttttttttct	tataaaactt	aaaaccttag	aggctattta	10560
agttgctgat	ttatattaat	tttattgttc	aaacatgaga	gcttagtacg	tgaacatga	10620
gagcttagta	cgtagccat	gagagcttag	tacgttagcc	atgagggttt	agttcggtta	10680
acatgagagc	ttagtacgtt	aaacatgaga	gcttagtacg	tgaacatga	gagcttagta	10740
cgtactatca	acaggttgaa	ctgcggatct	tgcgcccgca	aaaattaaaa	atgaagtttt	10800
gacggtatcg	aacccagag	tcccgtctcag	aagaactcgt	caagaaggcg	atagaaggcg	10860
atgcgctgcg	aatcgggagc	ggcgataccg	taaagcacga	ggaagcggtc	agccattcgc	10920
ccgccaagct	cttcagcaat	atcacgggta	gccaacgcta	tgctctgata	gcggtccgcc	10980
acacccagcc	ggccacagtc	gatgaatcca	gaaaagcggc	cattttccac	catgatattc	11040
ggcaagcagg	catcgccatg	ggtcacgacg	agatcctcgc	cgtcgggcat	ccgcgccttg	11100
agcctggcga	acagttcggc	tggcgcgagc	ccctgatgct	cttcgtccag	atcatcctga	11160
tcgacaagac	cggcttccat	ccgagtacgt	gctcgctcga	tgcgatgttt	cgcttggtgg	11220
tcgaatgggc	aggtagccgg	atcaagcgta	tgcagccgcc	gcattgcac	agccatgatg	11280
gatactttct	cggcaggagc	aaggtgagat	gacaggagat	cctgccccgg	cacttcgccc	11340
aatagcagcc	agtccttcc	cgcttcagtg	acaacgtcga	gcacagctgc	gcaaggaacg	11400
cccgtcgtgg	ccagccacga	tagccgcgct	gcctcgtctt	ggagttcatt	cagggcaccg	11460
gacaggtcgg	tcttgacaaa	aagaaccggg	cgccctgcg	ctgacagccg	gaacacggcg	11520
gcatcagagc	agccgattgt	ctgttgtgcc	cagtcatagc	cgaatagcct	ctccacccaa	11580
gcggccggag	aacctgcgtg	caatccatct	tgttcaatca	tgcgaaacga	tcctcatcct	11640
gtctcttgat	ccactagatt	attgaagcat	ttatcagggt	tattgtctca	tgagcggata	11700
catatttgaa	tgtattttaga	aaaataaaca	aatagggtt	ccgcgcacat	ttccccgaaa	11760
agtgccacct	gc					11772